

JOHN R. MATTINGLY*
DANIEL J. STANGER
SHRINATH MALUR*

GENE W. STOCKMAN
Of Counsel

JEFFREY M. KETCHUM
SCOTT W. BRICKNER
Registered Patent Agents

* Bar Membership Other Than Virginia

MATTINGLY, STANGER & MALUR, P.C.

ATTORNEYS AT LAW
104 EAST HUME AVENUE
ALEXANDRIA, VIRGINIA 22301

(703) 684-1120

PATENT, TRADEMARK
AND COPYRIGHT LAW
FACSIMILE: (703) 684-1157

Date: August 22, 2000

Attorney Docket No. ASA-911

To: Assistant Commissioner for Patents
Washington, D.C. 20231

Sir: Transmitted herewith for filing is the patent application of:

Inventor: M. KITAMURA et al (See Attached List)

For:

A COMPUTER SYSTEM AND A METHOD OF ASSIGNING
A STORAGE DEVICE TO A COMPUTER

Enclosed are:



23 Sheets of Drawings



This application is being filed without an executed Declaration.



Priority is claimed from Japanese Application No. 11-241024
filed August 27, 1999. ☐ A certified copy is attached herewith.



Copies of the disclosure documents listed on the attached PTO 1449 form and
☒ discussed in the specification or ☐ attached Information Disclosure Statement.



A verified statement to establish small entity status under 37 CFR 1.9 and 1.27.



Specification: Abstract ☒, Description 40 pages; and 16 claim(s).



Preliminary Amendment.



Executed Declaration.

The filing fee is calculated as shown below:

Small Entity

Large Entity

For:	No. Filed	No. Extra
Basic Fee		
Total Claims	16 - 20 = *	0
Indep Claims	4 - 3 = *	1
<input type="checkbox"/> Multiple Dependent Claim (s)		

* If difference is less than zero
then enter '0' in second column

Rate	Fee
	\$ 345
x 9	\$
x 39	\$
+ 130	\$
Total	\$

OR

Rate	Fee
	\$ 690
x 18	\$ 0
x 78	\$ 78
+ 260	\$ 0
Total	\$ 768



A check in the amount of \$ 768.00 is enclosed for the filing fee.



The Commissioner is hereby authorized to charge any additional fees that may be required to
Deposit Account No. 50-1417.

Respectfully Submitted,

By:

John R. Mattingly

Registration No. 30,293

- 1 -

A COMPUTER SYSTEM AND A METHOD OF ASSIGNING
A STORAGE DEVICE TO A COMPUTER

BACKGROUND OF THE INVENTION

The present invention relates to a computer system and a method for assigning a storage device to the computer system and more particularly, to a method
5 for assigning a storage device to a computer in a computer system including a storage subsystem having a fibre channel interface.

As the amount of information treated in a computer system for use in companies, corporations, etc.
10 is drastically increased, the capacity of a storage device such as a disk for storage of data has been increased steadily in these years. For example, a magnetic disk storage system having a capacity of the order of terabytes is very common. With respect to such
15 a disk storage system, there is a technique by which a single storage device subsystem is made up of a plurality of types of logical disks (which will be sometimes referred to merely as disks), e.g., as disclosed in U.S. Patent No. 5,956,750. Disclosed in
20 the disclosure is, more specifically, a disk subsystem which is made up of disks having different RAID levels such as RAID5 and RAID1 as devices (logical disks) to be accessed by a host computer, or made up of disks having different access rates as actual magnetic disks
25 (physical disks) of logical disks. A user can

selectively use the devices according to the access rates of the respective devices.

Appearance of a fibre channel technique as an interface between a host computer and a peripheral device such as a disk has also led to the fact that a plurality of host computers and a plurality of storage devices are connected by a single fibre channel cable to form a computer system. In such a computer system, each of the host computers can directly access any of the storage devices on the fibre channel. For this reason, the computer system can be expected to share data with the respective host computers and to reduce the load of a network, when compared with the prior art wherein each of the host computers has a storage device.

SUMMARY OF THE INVENTION

In the aforementioned prior art, the numbers and types of devices accessible by each host computer can be remarkably increased. However, as the numbers and types of devices accessible by each host computer is increased, it has been difficult to control the devices by each host computer. This system is advantageous in that many devices can be accessed by a single host computer, but is disadvantageous in that it is difficult for the user to select suitable one of the devices which is to be used for a given business. In particular, in the case of a computer system wherein the host computers and devices connected by fibre channels, one of the host

computers can access even one of the devices which is not initially intended to be used by the host computer in question. For this reason, the host computer may make unauthorized access to the device being used by the other host computer, thus leading to data destruction of the device.

In order to solve such a problem, there is disclosed in JP-A-10-333839 a method by which storage devices connected by fibre channels can be accessed only by a specific host computer. However, when a plurality of storage devices (or devices) are used or when different types of devices are mixedly used, its processing becomes still complex and each host computer must always take the types of the devices into consideration.

It is therefore an object of the present invention to facilitate setting of devices and assignment of the devices to host computers such that each of the host computers can use necessary one of the devices confirming to its application at a required time for a required time.

In a preferred embodiment of the present invention, a computer system has a plurality of computers and a storage device subsystem connected to the plurality of computers. The storage device subsystem, which has a plurality of storage devices and a plurality of interfaces, is connected to the computers. One of the computers has a management means for holding

09542817 082200

therein data about the storage devices within the storage device subsystem and about a connection relationship between the computers and storage device subsystem. Each computer, when wishes to form a new
5 device, informs the management means of its capacity and type. The management means, when informed by the computer, selects one of the storage devices satisfying its request. And the management means instructs the storage device subsystem to set necessary data in such a
10 manner that the request computer can access the selected device. The management means also returns necessary data to the computer as a device assignment demander, such that the computer as the requester can modify its setting based on the data and can use the assigned
15 device.

In another embodiment of the present invention, a plurality of computers and a plurality of storage device subsystems are interconnected in a network. Any one of the computers has a management
20 means for holding therein information of the storage devices possessed by the storage device subsystems and a connection relationship between the computers and the storage device subsystems. Each storage device subsystem has a control means for allowing access from
25 the management means to the specified computer. Each computer, when requiring a new storage device, informs the management means of its capacity and type. The management means, when informed by the computer,

002220 " 082200

instructs the storage device subsystem to allow access from the associated computer thereto, whereby the computer can select one of the devices satisfying the request and the computer in question can access the storage device subsystem. The management means also returns predetermined data to the computer as a device assignment request computer. The computer as the device assignment request computer, on the basis of the data returned from the management means, modifies setting of the computer in such a manner that the computer can use the device assigned thereto.

Other objects and advantages of the present invention will become clear as the following description of the invention advances as detailed with reference to preferred embodiments of the invention as shown in accompanying drawings.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a block diagram showing an exemplary arrangement of a computer system in accordance with a first embodiment of the present invention;

Fig. 2 shows an exemplary structure of a logical device management table held in a storage device subsystem;

Fig. 3 shows an exemplary structure of a host computer management table held by a control manager;

Fig. 4 is a flowchart showing a flow of operations implemented by a volume manager of a host

computer;

Fig. 5 is a flowchart showing a flow of operations implemented by the control manager;

Fig. 6 is a flowchart showing a flow of operations implemented by a volume manager in a device returning process;

Fig. 7 is a flowchart showing a flow of operations implemented by the control manager in the device returning process;

Fig. 8 is a block diagram showing an exemplary arrangement of a computer system in accordance with a second embodiment of the present invention;

Fig. 9 shows an exemplary structure of a logical device management table held by a control manager;

Fig. 10 shows an exemplary structure of a table for management of a correspondence relationship between host computers and WWNs held by the control manager;

Fig. 11 is a flowchart showing a flow of operations implemented by the control manager;

Fig. 12 is a diagram for explaining a zoning function of a fibre channel switch;

Fig. 13 is a block diagram of an exemplary arrangement of a computer system in accordance with a third embodiment of the present invention;

Fig. 14 is a flowchart showing a flow of operations by a client program;

09642617 082200

Fig. 15 is a flowchart showing a flow of operations by a control manager of a file server;

Fig. 16 is a flowchart showing a flow of operations by a server program of file server;

5 Fig. 17 is a flowchart showing a flow of operations implemented by the control manager when a file system is expanded;

Fig. 18 is a block diagram of an exemplary arrangement of a computer system in accordance with a
10 fourth embodiment of the present invention;

Fig. 19 is a flowchart showing a flow of operations by a volume manager;

Fig. 20 shows an exemplary structure of an LVOL management table;

15 Fig. 21 is a flowchart showing a flow of operations implemented by a volume manager when an LVOL is expanded;

Fig. 22 is an exemplary structure of a cluster data table; and

20 Fig. 23 is a flowchart showing a flow of operations implemented by a control manager when the LVOL is expanded.

DESCRIPTION OF THE EMBODIMENTS

25 Fig. 1 is a schematic block diagram showing an exemplary arrangement of a computer system in accordance with an embodiment of the present invention. The computer system includes a plurality of host computers

002280 2124950

1a, 1b, ..., and 1n (which will be generally denoted by a host computer 1), a storage device subsystem 2 connected to the host computer 1, a managing host computer 3, a network 4, and a substorage device 5.

5 Each of the host computers 1a, 1b, ..., which has a CPU, a memory and so on, can perform a predetermined function by its CPU which reads out an operating system and an application program from the memory and executing them.

10 The storage device subsystem 2 has a plurality of disk units 21, a disk controller 22, a plurality of ports 23 connected to the host computer 1, an interface 24 for connection with the substorage device 5, and a network interface 25 for connection with the network 4.

15 The storage device subsystem 2 in the present embodiment causes the host computer 1 to look as if there were a single or a plurality of logical devices by combining the plurality of disk units 21. Of course, the storage device subsystem 2 may cause the host computer 1 to look

20 as if there were one logical device for each disk unit 21.

 The port 23 may comprise, for example, an interface such as SCSI (Small Computer System Interface) when the host computer 1 to be connected therewith is a

25 computer in a so-called open system. When the host computer 1 is a so-called mainframe, on the other hand, the port 23 may comprise a channel interface such as ESCON (Enterprise System CONnection). The ports 23 may

002220 " 0822460

be an identical type interfaces or may include interfaces of different types. In the present embodiment, explanation will be made in connection with a case where the ports 23 are all an identical SCSI type of
5 interfaces.

The disk controller 22 has a processor 221, a cache memory 222, and a control memory 223. The processor 221 controls access from the host computer 1 to the disk units 21. In particular, when the storage
10 device subsystem 2 causes the host computer 1 to look as if the subsystem had not the single disk units 21 but a single or a plurality of logical devices corresponding to the single disk unit or a combination of the plurality of disk units as in a disk array, the
15 processor 221 performs its operation and management. The disk controller 22 communicates with the managing host computer 3 via the network interface 25.

In order to increase an access processing rate from the host computer 1, the cache memory 222 stores
20 data frequently read out from the host computer 1 therein or temporarily stores write data from the host computer 1 therein. When a part of the cache memory 222 is pretended as a single or a plurality of logical devices, the memory part can also be used as a device
25 which can eliminate access to the magnetic disk unit.

The control memory 223 is used to store therein a program to be executed by the processor 221 or to store therein data for management of the logical

09642817 082200

device or devices formed as a combination or combinations of the disk units 21.

Provided in each of the host computers 1a, 1b,... is software called a volume manager 11. The volume manager 11 operates to communicate with a control manager 31 provided in the managing host computer 3. Each host computer 1 has an interface (I/F) 12 and is connected at its interface 12 with the associated port 23 of the storage device subsystem 2.

Explanation will next be made as to how to manage the logical devices within the storage device subsystem 2.

As has been explained earlier, the storage device subsystem 2 causes the host computer 1 to look as if the storage device subsystem 2 had a single or a plurality of logical devices corresponding to a combination or combinations of the disk units 21 or had logical devices each corresponding to each of the disk units 21. The storage device subsystem 2 also causes the host computer 1 to look as if a part of the cache memory 222 of the subsystem like a single or a plurality of logical devices. The number of disk units 21 in the storage device subsystem 2 has no correlation with the number of logical devices.

Fig. 2 is an exemplary structure of a logical device management table for holding data for management of logical devices by the storage device subsystem 2. The logical device management table holds a relationship

002280 7 082200 09542817

of a set of items of size 62, configuration 63, state 64, path 65, target ID 66, and LUN 67 to a logical device number 61. Set in the item 'size' 62 is data indicative of the capacity of a logical device specified by the logical device number 61. The item 'configuration' 63 indicates data indicative of the configuration of the logical device, for example, RAID (Redundant Arrays of Inexpensive Disks) of the disk units 21. When the disk units are assigned to logical devices, the configuration item 63 sets therein data indicative of the type of the RAID such as RAID1 or RAID5. When a part of the cache memory 222 is assigned as the logical devices, the configuration item 63 sets 'cache' therein; while, when the single disk unit is assigned, the item 63 sets data indicative of the state of 'single disk unit' therein. The item 'state' 64 sets data indicative of the state of the logical devices therein. The state item has 'online', 'offline', 'not mounted', and 'offline trouble'. The 'online' indicates a state wherein the logical devices are in their normal operation and can be accessed by the host computer 1. The 'offline' indicates a state wherein the logical devices are defined and are in their normal operation, but cannot be accessed by the host computer 1. This state corresponds to a case where the logical devices have been so far used by the host computer 1 became unused because it became unused for the host computer 1 to use them. The 'not mounted' indicates a

state wherein the logical devices are not defined and cannot be accessed by the host computer. The 'offline trouble' indicates a state wherein a trouble takes place in the logical devices and thus the devices cannot be
5 accessed by the host computer. Set in the item 'path' 65 is data indicative of one of the ports 23 connected with associated one of the logical devices. A unique number in the storage device subsystem 2 is assigned to each of the ports 23, and the numbers of the ports 23
10 connected with the associated logical devices are recorded in the column 'path'. The target ID 66 and LUN 67 are each an identifier for discrimination between the logical devices. In this example, the SCSI-Ids and LUNs used when the devices are accessed by the host
15 computer 1 on the SCSIs are used as these identifiers.

One logical device can be connected to a plurality of ports and a plurality of host computers 1 can access an identical logical device. In this case, the logical device management table has a plurality of
20 entries relating to the logical device created therein. For example, in such a logical device management table as shown in Fig. 2, the device having a logical device number of 2 is connected to two ports 23 having port numbers 0 and 1. For this reason, two items of the
25 logical device numbers of 2 are present. When a single logical device can be accessed by the plurality of ports 23 in this way, the target ID and LUNs associated with the respective paths 65 do not have to be the same

09642817 082200

respectively and may be different as shown in Fig. 2.

Data described in the logical device management table are sent via the interface 24 to the managing host

computer 3 at suitable timing or as when a trouble

5 occurred in the storage device subsystem 2 to change the configuration. Thus, the managing host computer 3 also holds a logical device management table similar to the table shown in Fig. 2.

Fig. 3 is an exemplary structure of a host
10 computer management table held by the control manager 31 of the managing host computer 3.

The host computer management table holds management data as a set of a host computer name 71, a port number 72, an interface number 73 and a logical
15 device number 74 in order for the managing host computer 3 to manage assignment of devices to the host computers 1.

The port number 72 and logical device number 74 are numbers defined in the storage device subsystem 2
20 and are data for discrimination between the ports 23 and logical devices of the storage device subsystem 2. Set in the items of 'port number' 72 and 'logical device number' 74 are the number assigned to a port connected with the host computer 1 identified by an identifier set
25 in the host computer name item 71 as well as the number or numbers assigned to logical device or devices assigned to the host computer respectively. The interface number 73 is used to manage the interface 12

of each host computer 1. The interface number 73 becomes necessary, in particular, when the single host computer 1 has a plurality of interfaces 12. A set of the port number 72 and interface number 73 is an
5 important factor to show a connection relationship between the host computer 1 and associated logical device or devices. For example, the host computer 1b shown in Fig. 1 has two interfaces 12 which are connected to different ports 23 respectively. In such a
10 case, even when one of such interfaces or a line connected between one interface and storage device subsystem 2 cannot be used due to occurrence of a defect therein, its processing can be continued so long as the host computer is connected at the other interface with
15 the logical device, thus enhancing its reliability.

The managing host computer 3 assigns the logical devices to the host computers 1, by referring to its own host computer management table and the logical device management table received from the storage device
20 subsystem 2. Explanation will then be made as to how to assign the devices.

Fig. 4 is a flowchart showing a flow of operations implemented by the volume manager 11 of each host computer 1. The operations are carried out when a
25 user of the host computer 1 or an application program to be run on the host computer 1 requires new device or devices.

At a step 1001, the volume manager 11 obtains

data on the number and type of the devices required by the user or application program. The user or application program specifies data including its capacity, performance conditions and reliability level as the device data. The term "the capacity of the device" refers to the size of the device as already explained above. The user can specify, as the performance conditions, performance data of device access rate, etc. such as, for example, low-speed disk drive high-speed disk drive, or cache residence disk drive. The user can specify, as the reliability level, device reliability data such as, for example, RAID0, RAID1, RAID5, double path or remote mirror. For the double path, when the host computer 1 has a plurality of interfaces, a plurality of paths are provided so that the host computer 1 can access an identical device utilizing the plural interfaces. For the double path, even when one of the paths cannot be used for some reason, the host computer 1 can access the device with use of the other path. The "remote mirror" means that a copy of the devices in the storage device subsystem 2 is provided to the substorage device 5. Thus, even when the storage device subsystem 2 itself cannot be operated due to factors such as earthquake or fire, the device data are held in the substorage device 5, thus enabling increase of its reliability.

At a next step 1002, the volume manager 11 searches for a set of target ID and LUN not used on the

interface 12 of the host computer 1 in question.

At a step 1003, the volume manager 11 transmits the capacity, performance conditions and reliability level specified at the step 1001 as well as
5 the set of target ID and LUN not used and searched at the step 1002 to the control manager 31 of the managing host computer 3 to request assignment of a new device. The control manager 31 searches for a device to be assigned on the basis of the received data and returns
10 data to specify the host computer interface number, target ID and LUN to be used for device accessing. How to operate the control manager 31 in this case will be explained later.

At a step 1004, the volume manager 11 receives
15 the data from the control manager 31. At a step 1005, the volume manager 11 modifies setting of the host computer 1 on the basis of the data received from the control processor 31 in such a manner that the host computer 1 can use a new device. In the case of a so-
20 called open operating system, the host computer 1 makes access to each device, so that a device file is prepared for each device and its access is carried out for the device file. The device file is usually prepared when the host computer 1 underwent a device configuration
25 operation, and no device file is created for the device which was not present during the device configuration operation. For this reason, at the step 1004, a device file for a newly assigned device is created. More

002220 182450

specifically, in the case of a Solaris operating system for example, a command 'drvconfig' or 'drives' is used to recognize the new device and create the device file, whereby the host computer 1 can access the newly assigned device.

At a final step 1006, the volume manager 11 informs the user or application program of data about the assigned device file name, target ID and LUN, thus terminating its operation.

Fig. 5 is a flowchart showing a flow of operations carried out by the control manager 31 of the managing host computer 3 during assignment of the new device.

At a step 1101, the control manager 31, when receiving the data about the device size, performance conditions and reliability level from the host computer 1, searches for the device satisfying the request by referring to its own logical device management table and host computer management table. In this example, the device to be searched for is denoted in the state item 64 of the logical device management table as the 'offline' (step 1102). The control manager 31, on the basis of its search result, judges the presence or absence of the device of the 'offline' state satisfying the request (step 1103).

When finding the device of the 'offline' state satisfying the request, the control manager 31 determines a port number, target ID and LUN for connection of

00222200 272224550

the device to the host computer 1, on the basis of the target ID and LUN received from the host computer 1 as well as data set in the logical device management table and host computer management table (step 1104).

5 Subsequently the control manager 31 sets the device of the logical device number found at the step 1103 to allow the host computer 1 to be able to access the device with the port number, target ID and LUN determined at the step 1104, and instructs the storage
10 device subsystem 2 to put its state in 'online'. The storage device subsystem 2 performs setting operations according to the instruction from the control manager 31, and returns its result to the control manager 31 (step 1105).

15 The control manager 31, when receiving the result from the storage device subsystem 2 (step 1106), returns the interface number, target ID and LUN to the volume manager 11 of the host computer 1 as the request demander (step 1107).

20 When failing to find a device of the 'offline' state satisfying the request at the step 1103, on the other hand, the control manager 31 searches for the presence of a logical device number having the 'not mounted' set in state item 64 of the logical device
25 management table (step 1108). The presence of a logical device number having the 'not mounted' causes the control manager 31 to inform of the storage device subsystem 2 of data about the device size, performance

002220" / FEB 4 1960

5

10

15

20

25

operation (step 1204).

Fig. 7 is a flowchart showing a flow of operations to be carried out by the control manager 31 in the unnecessary-device returning process of the host computer 1.

The control manager 31 receives the interface number, target ID and LUN from the host computer 1 (step 1301). The control manager 31, on the basis of the received interface number, target ID and LUN, instructs the storage device subsystem 2 to put the device to be returned in its offline state. In response to the instruction, the storage device subsystem 2 puts the specified device in the offline state, and returns the logical device management table reflecting its result to the control manager 31 (step 1302). The control manager 31, when receiving the logical device management table from the storage device subsystem 2, holds it therein and completes its operation (step 1303).

Although the managing host computer is provided and the control manager is positioned therein in the foregoing first embodiment, it is not necessarily required to provide the function of the control manager in the managing host computer. For example, the control manager function may be provided in any of the host computers, 1a, 1b, ... The control manager function also may be provided in the storage device subsystem. In this case, the host computers 1a, 2b, ... sends a request and receives its data directly to and from the

storage device subsystem via their interfaces.

Fig. 8 schematically shows an arrangement of a computer system in accordance with a second embodiment of the present invention. The computer system of the present embodiment includes a plurality of host computers 1 (host computers 1a, 2b, ..., and 1n), a plurality of storage device subsystems 2a, 2b, ... and 2m, a managing host computer 3, a network 4, and a fibre channel switch 6.

The host computer 1, as in the first embodiment, has a volume manager 11. The volume manager 11 operates to communicate with the control manager 31 provided in the managing host computer 3. Further, each of the host computers 1 has an interface (I/F) 12 which in turn is connected to a fibre channel switch 8 via its interface 12.

Similarly to the storage device subsystem 2 in the first embodiment, each of the storage device subsystems 2a, 2b, ... and 2m includes a disk unit 21, a disk controller 22, a port 23, and a network interface (network I/F) 25 connected to the network. Although a plurality of such disk units 21 and a plurality of such ports 23 may be provided as in the first embodiment, explanation will be made below in connection with the single disk unit and single port, for simplicity.

The fibre channel switch 8 has a plurality of ports 81 which in turn are connected to the associated interfaces 12 of the host computers 1a, 1b, ... and to

connecting host computer name 107 for each device. The size 103, configuration 104, state 105, LUN 106 are the same as those in the logical device management table in the first embodiment. The WWN 102 is data which is set
5 in the port 23 of the storage device subsystem 2 and which is uniquely assigned to each fibre channel interface for identification of each port. The WWN 102 is also called N_PORT_NAME. The connection host computer name 107 is used to identify the host computer
10 which is allowed to be connected to the device in question.

Basically, when the plurality of host computers 1 connected to the fibre channel switch 8 can freely access any of the storage device subsystems 2, it
15 may, in some cases, become a problem from the viewpoint of system security. In order to solve the problem with such system security, for example, JP-A-10-333839 discloses a technique in which only a specific host computer can access a storage device connected therewith
20 by a fibre channel. It is assumed even in the present embodiment that, in order to maintain the system security, the storage device subsystem 2 has such a means for maintaining the security as disclosed in JP-A-333839. However, this is not associated directly with
25 the essence of the present invention and thus detailed explanation thereof will be omitted.

In the present embodiment, the WWN 109 is provided even to the interface 12 of each host computer

1. The managing host computer 3, on the basis of such a table as shown in Fig. 10, manages a set of a host computer name 108 and WWN 109.

Explanation will then be made as to the
5 operation of the volume manager 11 and control manager 31.

In the present embodiment, the operation of the volume manager 11 when a new device is assigned to the host computer is basically the same as that in the
10 first embodiment of Fig. 4. More specifically, the volume manager 11, when receiving data on the number and type of the necessary device from the user or application program, requires the control manager 31 to assign the new device thereto on the basis of the received
15 data. After the control manager 31 finishes the assignment of the new device, the volume manager 11 modifies setting of the device such that the host computer 1 can use the new device.

Fig. 11 shows a flowchart showing a flow of
20 operations carried out by the control manager 31 at the time of the assignment of the new device in the present embodiment. The operations carried out by the control manager 31 are substantially the same as those by the control manager in the first embodiment of Fig. 5. In
25 this connection, parts having substantially the same functions as those in Fig. 5 are denoted by the same reference numbers in Fig. 11. Explanation will be mainly made as to only parts having functions different

from those in Fig. 5, and explanation of the parts having the same functions as those in Fig. 5 is omitted.

In the present embodiment, the storage device subsystem 2 disables, in its initial state, access from
5 all the host computers 1 to prevent the device from being accessed by the not-assigned host computers. To this end, the control manager 31, when instructing the storage device subsystem 2 to put the device in its online state at a step 1105, also instructs the storage
10 device subsystem 2 to enable access from the host computer 1 to the device to be newly assigned. In this instruction, the control manager 31 informs the storage device subsystem 2 of the WWN of the host computer 1 to enable the access to the device. The storage device
15 subsystem 2, when the host computer 1 wants to access the device, judges permission or non-permission of the access on the basis of the WWN received from the control manager 31 (step 2105).

After the operation of the step 2105, the
20 control manager 31 modifies the setting of the fibre channel switch 8. Consider, for example, a case where , as shown in Fig. 12, host computers A and B can access disk units (devices) **a** and **b**, while a host computer C can access only a disk unit (device) **c**. In this case,
25 the control manager 31 instructs the fibre channel switch 8 to modify its path setting such that the host computer C cannot access ports **d** and **e** connected to the disk units **a** and **b** from the port **c** connected to the host

002220 "CTB4500
0564281" 082200

computer C. It can be arranged to look as if there were two switches. Such path setting is called zoning. The zoning can prevent a host computer from accessing a device which should not be accessed from the host computer. Further, since data flows are separated, its performance can be improved (step 2106).

After the above operation, the control manager 31 performs the operations of the steps 1106 and 1107.

Fig. 13 schematically shows a block diagram of an exemplary arrangement of a computer system in accordance with a third embodiment of the present invention. In the computer system of the present embodiment, a plurality of host computers 1a', 1b', ..., and 1n' (which are generally referred to as host computer 1') are connected to a file server 9 via interfaces (I/F's) 12' and a network 4. The file server 9 is connected at its interfaces 92(I/F) 92 to the storage device subsystem 2. A storage device subsystem 2 and substorage device 5 as a storage device located at a remote place are substantially the same as those in the first embodiment.

The file server 9 has a network interface 91 connected to the host computers 1', a plurality of interfaces 92 connected to the storage device subsystem 2, a control manager 93 and a server program 94. The control manager 93 performs device assigning operation as required as in the control manager 31 of the first embodiment. The server program 94 is a file server

program which offers file access via the network such as NFS (Network File System). The server program 94 offers a means for allowing the host computers 1' to be able to access a file system prepared in the storage device subsystem 2 by the file server 9. Such an NAS (Network Attached Storage) arrangement may be possible that the storage device subsystem 2 and file server 9 cause the host computer 1' to look as if there were a single storage device.

10 A client program 11' of each of the host computers 1' functions to communicate with the server program 94 of the file server 9 in such a manner that an application program running on the host computer 1' can use the file system prepared in the storage device subsystem 2 by the file server 9. The client program 11' may be built in an operating system (not shown) of the host computer 1' depending on the system configuration. The client program 11' requires the control manager 93 to make a new file system or to
15 modify the size of the existing file system. In order that the host computer 1' can modify the size of the existing file system during run of the computer 1', the storage device subsystem 2 of the present embodiment has a function of moving the data present in a logical
20 device to a physical disk unit different from a physical disk units having the logical device already formed therein. As a specific technique for realizing such a function, such a known technique as, e.g., in U.S.

002220" 182450

Patent No. 5,956,750 can be employed. Therefore detailed explanation thereof will be omitted in this specification.

Fig. 14 is a flowchart showing a flow of operations carried out when the client program 11' of the host computer 1' forms a new file system. The operations are carried out when a user using the host computer 1' or an application program running on the host computer 1' requires a new file system area. The client program 11' accepts specification of data about a necessary device according to a request from the user or application program. The accepted data include, as in the step 1001 in the first embodiment of Fig. 4, a capacity, performance conditions and reliability level of the necessary device (step 2001). The client program 11' then transmits the specified capacity, performance conditions and reliability level specified at the step 2001 to the control manager 93 to require the new file system area. The control manager 93, on the basis of the data accepted from the client program 11', searches for and prepares the assignable device area and returns its result to the client program 11'. The then operation of the control manager 93 will be explained later (step 2002). The client program 11' receives a response to the request of the new area from the control manager 93. The response then received includes a mount point such as a host name or host IP address of the file server in the case of NFS, and a directory name (step

2003). The client program 11', on the basis of the data received from the control manager 93, mounts the file system (step 2004). Finally, the client program 11' informs the user or application program of the assigned
5 mount point and terminates its operation (step 2005).

Fig. 15 is a flowchart showing a flow of operations carried out by the control manager 93 in response to the request of the new area from the client program 11'. The operations are carried out basically
10 in the same manner as in those of the control manager 31 in the first embodiment of Fig. 5. In this connection, the operation of the step 1107 in Fig. 5 is changed to the operations of steps 2107, 2111 and 2112. Although such data as target ID is passed to the host computer as
15 the request originator at the step 1107 in Fig. 5, such data is processed in the present embodiment. To this end, the control manager 93 passes data about the device such as the target ID to the server program 94 (step 2107), and receives data about the mount point from the
20 server program 94 (step 2111). And the control manager 93 passes the mount point data received from the server program 94 to the client program 11' and then terminates its operation (step 2112).

Fig. 16 is a flowchart showing a flow of
25 operations carried out by the server program when receiving the device data from the control manager. The server program 94, when receiving the device data from the control manager 93 (step 2201), performs device

reconfiguring operation of the file server 9. More in detail, this operation is substantially the same as that of the step 1005 in the first embodiment of Fig. 4 (step 2202). Subsequently the server program 94 creates a
5 file system in the new device (step 2203), and returns data indicative of the mount point of the file system to the control manager 93 (step 2204).

Through the above operation, there can be added a new file system which can be used by the host
10 computer 1'.

Fig. 17 is a flowchart showing a flow of operations carried out by the control manager 93 at the time of modifying the size of the existing file system. The operation of Fig. 17 is different from the operation
15 of Fig. 15 at the time of requiring a new file system in respects which follow.

When it is desired to modify the size of the existing file system, the user or application program issues a request to the client program 11' containing
20 data such as the mount point of the file system to be modified in its size and the size to be expanded or reduced. The client program 11', using the data specified by the user or application program, requires the control manager 93 to modify the size of the file
25 system. The control manager 93 receives the mount point of the file system to be processed and the size to be expanded, from the client program 11' (step 2301). The control manager 93, on the basis of the mount point

received from the client program 11', acquires data about the target ID, LUN, etc. of the logical device storing therein the file system to be processed, to identify the logical device. And the control manager 93
5 acquires the type of the logical device, that is, data on its reliability, performance, etc. (step 2302). Subsequently, as when a new file system is added, the control manager 93, on the basis of the data acquired at the steps 2301 and 2302, allocates the logical device of
10 the same type as the initial logical device and having the same capacity as the size of the file system after modified (steps 1102 to 1110). Thereafter control manager 93 instructs, at a step 2304, the storage device subsystem 2 to move or transfer the data from the
15 logical device having the file system so far recorded therein to the new secured logical device. The data transfer is carried out transparently from the server program 94. Since the host computer 1' makes access to the storage device subsystem 2 via the server program
20 94, this operation becomes transparent vent from the host computer 1'. Accordingly, it is unnecessary to stop the operation during the data transfer. After finishing the data transfer, the control manager 93 instructs the server program 94 to expand the file
25 system. Even when the capacity of the device is actually increased, all the capacity of the expanded file system cannot be used so long as the file system is not reformed. The control manager 93, after instructing

05642817 082200

the server program 94 to expand the file system, informs the client program 11' of completion of the operation and terminates its operation (step 2305).

Through the above operation, the size
5 modification of the existing file system can be realized while the host computer 1' is being run. When it is desired to modify the size of the existing file system, the client program 11', after receiving its notification from the control manager 93, can use the expanded file
10 system as it is. In this case, thus, it is unnecessary to carry out the operations of the steps 2004 and 2005 in Fig. 14.

Fig. 18 is a schematic block diagram showing an exemplary arrangement of a computer system in
15 accordance with a fourth embodiment of the present invention. The computer system of the present embodiment includes a plurality of host computers 1" (host computers 1a", 1b", ..., and 1n"), a managing host computer 3, a storage device subsystem 2', and a
20 substorage device 5. The host computers 1" and storage device subsystem 2' are connected each other by a fibre channel switch 8. The host computers 1", storage device subsystem 2' and fibre channel switch 8 are mutually connected by means of a network 4. The fibre channel
25 switch 8 having a plurality of ports 81 performs switching operation of connections between the ports 81 to realize data transfer to the devices connected thereto. The fibre channel switch 8 also has a network

002230 27824560

interface 82 for communication via the network 4. Each of the host computers 1" has a volume manager 11" and a single or a plurality of interfaces 12. The interface or interfaces 12 of the host computer 1" are connected to any of the plurality of ports 81 of the fibre channel switch 8.

The storage device subsystem 2' has an inter-controller connecting mechanism 27 which performs interconnection between a plurality of clusters 26.

Each of the clusters 26 has a channel processor 23', a drive processor 22', and a plurality of disk units 21. The channel processor 23' and drive processor 22' in the same cluster are coupled with each other by means of a bus 28 higher in data transfer rate than the inter-controller connecting mechanism 27. Each channel processor 23' having a single or a plurality of ports 231 is connected to the substorage device 5 or to the host computers 1" via the fibre channel switch 8. Each drive processor 22' is connected with a plurality of disk units 21. In the present embodiment, these disk units 21 are combined into a single or a plurality of logical devices, or the single disk unit 21 is made up of a single or a plurality of logical devices. In this connection, it is assumed that the single logical device cannot be made of any combination of the disk units 21 included in the plural clusters 26.

The channel processor 23' causes each host computer 1" to look as if there were a single or a

09542817 082200

plurality of logical devices, and accepts access from each host computer 1". On principle, the channel processor 23' manages logical devices formed by the disk units 21 belonging to the channel processor 23'. This is because communication between the channel processor 23' and drive processor 22' in the same cluster 26 can be made faster than the communication between the clusters. When the channel processor 23' in a given cluster 26 becomes inoperative due to a factor such as a trouble, however, the operation of the channel processor 23' is replaced by the channel processor 23' of another cluster 26. The channel processor 23' identifies the disk units 21 by which the logical device specified by the host computer 1" is formed, and passes its request to the right drive processor 22'. The drive processor 22' interprets the request received from the channel processor 23', generates disk access request to the disk unit 21 of which the logical device is composed, and sends the disk access request to each corresponding disk unit 21.

The host computer 1" has substantially the same structure as the host computer 1 in the first embodiment, but is slightly different therefrom in the function of the volume manager 11" running thereon. The volume manager 11" has a function of, in addition to the function of assigning and returning the logical device by the volume manager 11 in the first embodiment, combining a plurality of logical devices into another

logical device so as to cause the upper-level application program recognizes as if the plural logical devices were the single other logical device. The logical device formed by the volume manager 11" will be hereinafter denoted by LVOL to draw a distinction from the logical device managed by the storage device subsystem 2'. The volume manager 11" can combine a plurality of logical devices into a large single LVOL or divide a single logical device into a plurality of areas to cause the application program on the host computer 1" to use these areas as LVOLs. Further, the volume manager 11" can combine the existing LVOL and a new logical device to expand the capacity of the LVOL.

Fig. 19 is a flowchart showing a flow of operations carried out by the volume manager 11" when a volume is newly assigned in the present embodiment. The operations to be explained herein correspond to a step 1002' replaced by the step 1002, and steps 1005' and 1006' replaced by the step 1006 in the device assigning operation in the first embodiment of Fig. 4. The other steps are carried out in substantially the same manner as the corresponding steps in Fig. 4. Explanation will be made as to the operations of the steps 1002', 1005' and 1006'.

At the step 1002', the volume manager 11" searches an LVOL management table managed thereby for a set of WWN and LUN not used. An example of the LVOL management table is shown in Fig. 20. Registered in the

LVOL management table are data as a set of LVOL name
151, device file name 152, size 153, WWN 154 and LUN 155
for each device. The LVOL name 151 is an identifier for
distinction between LVOLs provided to the application
5 program by the volume manager 11". The device file name
152 is the name of the logical device forming an LVOL.
The volume manager 11" manages the logical devices
belonging to the LVOLs on the basis of the device file
name. The size 153 indicates the capacity of each
10 logical device forming an LVOL. Since one LVOL may form
a plurality of logical devices, a plurality of device
files may belong to one LVOL name.

At the step 1005', the volume manager 11"
creates a new LVOL using the logical device assigned by
15 the control manager 31, and registers its contents in
the LVOL management table. At the step 1006', the
volume manager 11" informs the user of the LVOL name
already assigned and terminates its operation.

Fig. 21 is a flowchart showing a flow of
20 operations carried out by the volume manager when the
capacity of the LVOL is expanded in response to a
request from the user or application program. When the
capacity of the LVOL is expanded, the volume manager
prepares a new logical device, and combine the prepared
25 logical device with the logical device forming the LVOL
to be expanded to thereby form a new LVOL. At this
time, it is common that the newly prepared logical
device of the same type as the logical device forming

002220" 082200
05642817

the LVOL to be expanded. In the present embodiment, the volume manager 11" judges the type of the logical device forming the LVOL to be expanded and allocates the same type of logical device.

5 In the operation, the volume manager 11" first receives data about the LVOL name of the LVOL to be expanded as well as the capacity to be expanded from the user or application program (step 2501). Next the volume manager 11" inquires the type of the logical
10 device forming the LVOL to be expanded of the control manager 31 (step 2502). The volume manager 11" searches the LVOL management table for a set of WWN and LUN not used (step 2503). The volume manager 11" sends to the control manager 31 the type of the logical device as
15 well as the set of WWN and LUN not used, acquired at the steps 2502 and 2503 (step 3504). The volume manager 11", when receiving the data on the newly assigned logical device from the control manager 31 (step 2505), performs reconfiguring operation of the host computer 1"
20 to allow the host computer 1" to be able to recognize the newly-assigned logical device (step 2506). Finally, the volume manager 11" adds the newly-assigned logical device to the LVOL to be expanded to expand the capacity of the LVOL and terminates its operation (step 2507).

25 When receiving a request about assignment of a new logical volume from the volume manager 11" at the step 1003 in Fig. 19 and at the step 2504 in Fig. 21, the control manager 31 searches for and assign the type

002220" 2784950

and capacity of the device required by the volume manager 11" in either case. To this end, the control manager 31 has, in addition to such a logical device management table as shown in Fig. 9, a cluster data
5 table having data about the clusters 26 in the storage device subsystem 2' set therein.

Fig. 22 is an exemplary structure of a cluster data management table. The cluster data management table sets therein entries for each cluster 26, that is,
10 a set of cluster number 161 for identification of the cluster, port number 162 of the port having the cluster, and WWN 163 assigned to the port, for each cluster 26. When a plurality of ports are provided to one cluster 26 as illustrated, the respective port numbers and WWNs are
15 set in the corresponding entries of the cluster in question. As already explained above, when a logical device was formed for the disk units 21 connected to the drive processor 22', it is desirable from the performance viewpoint that the port 231 in the same cluster can
20 access the logical device. The control manager 31, on the basis of the cluster data table, sets the devices in such a manner that the port 231 used for access from the host computer 1" and the drive processor 22' connected with the disk unit 21 formed by a logical device to be
25 newly assigned belong to the same cluster.

Fig. 23 is a flowchart showing a flow of operations in the device assigning process by the control manager 31. The device assigning operations of

09642817 082200

the present embodiment are carried out in substantially the same manner as those of the second embodiment of Fig. 11, but are partly different from the second embodiment due to a difference in structure between the storage device subsystems. More specifically, when a request to form a new device is issued from the control manager 31 to the storage device subsystem 2' at a step 1109, the storage device subsystem 2' forms a device in response to the request. When the new logical device is formed in the storage device subsystem 2', the control manager 31 receives data indicative of the cluster 26 in which the new logical device was formed therefrom (step 2610). The control manager 31 determines the port from which the new logical device can be accessed by referring to the data on the logical device received from the storage device subsystem 2' and by referring to the cluster data management table. The control manager 31, based on the unused LUN information, further determines the LUN of the device to be newly assigned (step 2604). The control manager 31 finally sends data of WWN, LUN, etc. necessary for access of the logical volume to be newly assigned to the volume manager 11' (step 2607).

Operations other than the operations already explained above are substantially the same as the operations of the control manager in the second embodiment of Fig. 11, and thus detailed explanation of parts having substantially the same function as those in

Fig. 11 is omitted.

In accordance with the foregoing embodiments, even when the host computer is operating, the device assigning operation can be carried out in response to its request. Further, even when a plurality of devices configured by the fibre channel switch are mixedly provided, the device satisfying a request from the host computer can be assigned to the host computer.

It should be appreciated that the present
10 intention is not to limit the invention only to the
foregoing embodiments shown but rather to include all
alterations, modifications and equivalent arrangements
possible within the gist and spirit of the invention.

WHAT IS CLAIMED IS:

1. A computer system comprising:

first computers;

a storage device subsystem including storage devices holding therein data to be accessed by said first computers; and

a second computer having therein device management data on said storage devices possessed by said storage device subsystem and also having host computer management data indicative of assignment states of said first computers to said storage devices,

wherein said first computer has request means for receiving a request of a new storage device from a user or application program and requiring said second computer to assign the new storage device,

said second computer has determination means, in response to the request from said request means, for determining one of the storage devices assignable to said first computer by referring to said device management data and said host computer management data and also has modification means for modifying setting of said storage device subsystem in such a manner that the storage device determined by said determination means can be accessed by said first computer.

2. A computer system as set forth in claim 1, wherein said storage device is at least part of storage areas of the physical storage devices provided in said storage device subsystem.

09542817 082200

4. A computer system as set forth in claim 3, wherein said determination means refers to said device management data and selects one of said storage devices which satisfies said conditions specified by said request means and which is in its offline state as an assignable storage device.

6. A computer system as set forth in claim 1, wherein said request means has means for receiving a request for modification of a capacity of an existing storage device and sending said request to said second computer, said second computer has, in response to the capacity modification request, means for selecting one of the storage devices which coincides in capacity with a capacity-modified storage device and also has means for instructing said storage device subsystem to move data held in said existing storage device to the storage device selected by said selection means, and said storage device subsystem has means for moving the data in response to said instruction.

7. A computer system as set forth in claim 1,

8. A computer system as set forth in claim 7, wherein said request means, in response to a request for expanding the capacity of the device already provided to said application program, issues a request for assignment of a new storage device having a capacity necessary for expansion of said expansion request to said second computer, and said device providing means for setting the newly-assigned storage device as part of said expansion-required device.

issuing a request for assignment of a new storage device to said second computer from said first computer;

determining in said second computer the storage device assignable to said first computer on the basis of said request;

instructing, in said second computer, said storage device subsystem to modify setting of the

13. A method as set forth in claim 9, further comprising the step of, in said first computer, combining said determined storage device and an existing

a storage device subsystem connected to said file server and having a plurality of storage devices therein

wherein each of said plurality of computers has request means for receiving a request of a new file area from a user or application program and requiring file server to assign the new file area, and

said file server, in response to the request from said request means, has means for determining one of the storage devices which has a storage area and to which said file can be set by referring to said device management data and said host computer management data, modification means for modifying setting of said storage device subsystem in such a manner that said file server can access the storage device determined by said determination means, setting means for modifying setting of said file server in such a manner that said determined storage device can be used, and file management means for creating said file area on said determined storage device.

16. A computer system as set forth in claim 15, wherein said file management means has means for transmitting data used for access to the created file area to the computer as the request sender, and said request sender computer has an application program which can access said created file are utilizing said data used for access.

002220" 082220

ABSTRACT OF THE DISCLOSURE

A computer system which has a plurality of computers and a storage device subsystem connected to the plurality of computers. The storage device subsystem has a plurality of storage devices and a plurality of interfaces, through which the subsystem is connected to the computers. One of the plurality of computers has a management means for holding therein data indicative of the storage devices and a connection relationship between the computers and storage device subsystem. Each computer, when wanting a new device, informs the management means of its capacity and type. The management means receives its notification and selects one of the storage devices which satisfies the request. And the management means instructs the storage device subsystem to set predetermined data in such a manner that the computer can access the selected device. The management means also returns predetermined data to the computer as a device assignment requester, the assignment requester computer modifies setting thereof to allow the computer in question can use the assigned device.

002220" 022450

FIG. 1

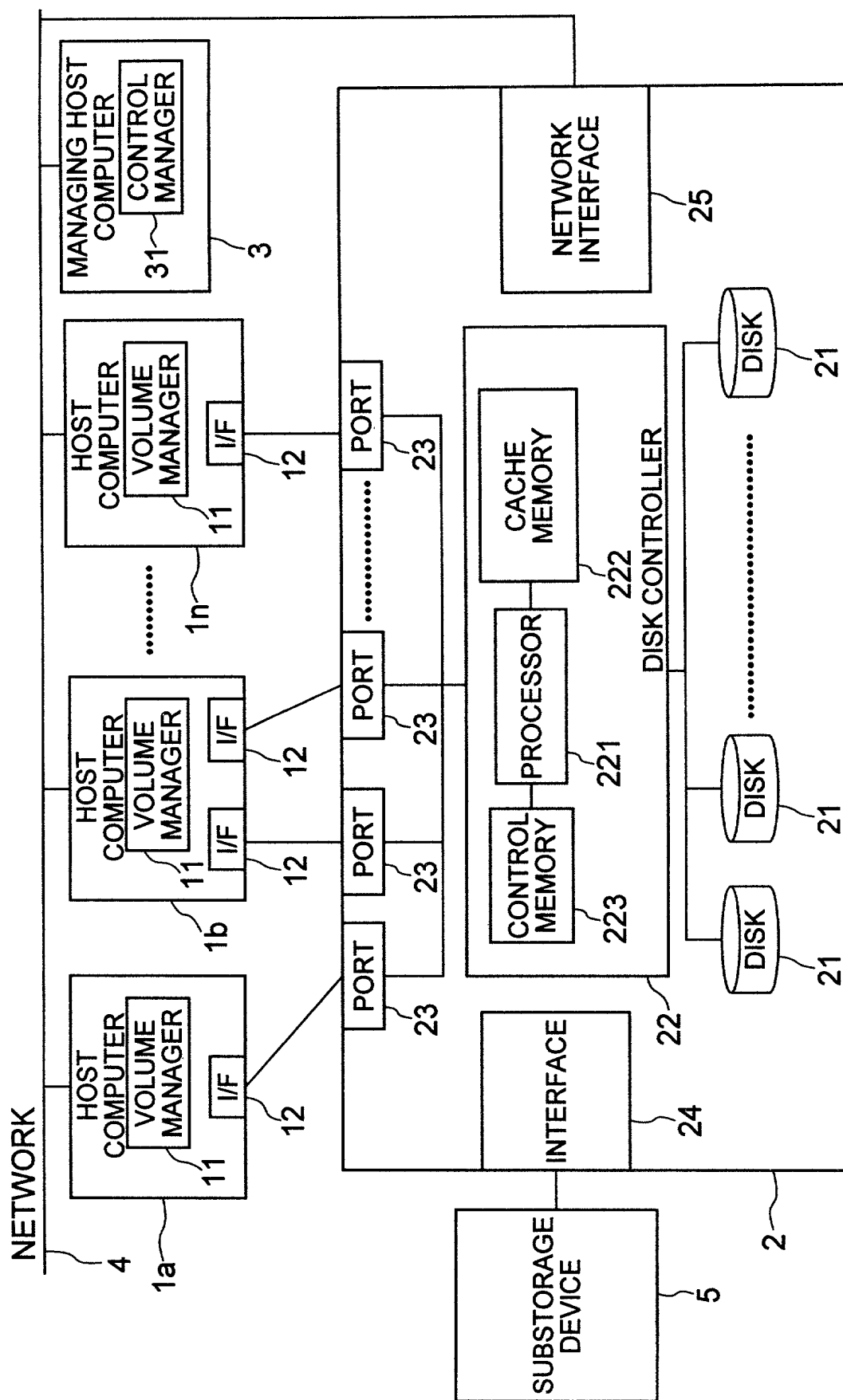


FIG. 2

61	62	63	64	65	66	67
LOGICAL DEVICE NUMBER	SIZE	CONFIGURATION	STATE	PATH	TARGET ID	LUN
0	1000000	RAID1	ONLINE	0	0	0
1	1000000	RAID5	ONLINE	-	-	-
2	35000	CACHE	ONLINE	0	0	2
2	35000	CACHE	ONLINE	1	0	0
.
.
.
.
.
k	-	-	NOT MOUNTED	-	-	-
k+1	1000000	SINGLE DISK	OFFLINE TROUBLE	0	1	0
k+2	1000000	RAID1	OFFLINE	-	-	-
n	-	-	NOT MOUNTED	-	-	-

FIG. 3

71 HOST COMPUTER NAME	72 PORT NUMBER	73 INTERFACE NUMBER	74 LOGICAL DEVICE NUMBER
HOST 1	0	0	0,1
HOST 2	1	0	1
HOST 2	2	1	3
⋮	⋮		
HOST N	0	0	n

FIG. 4

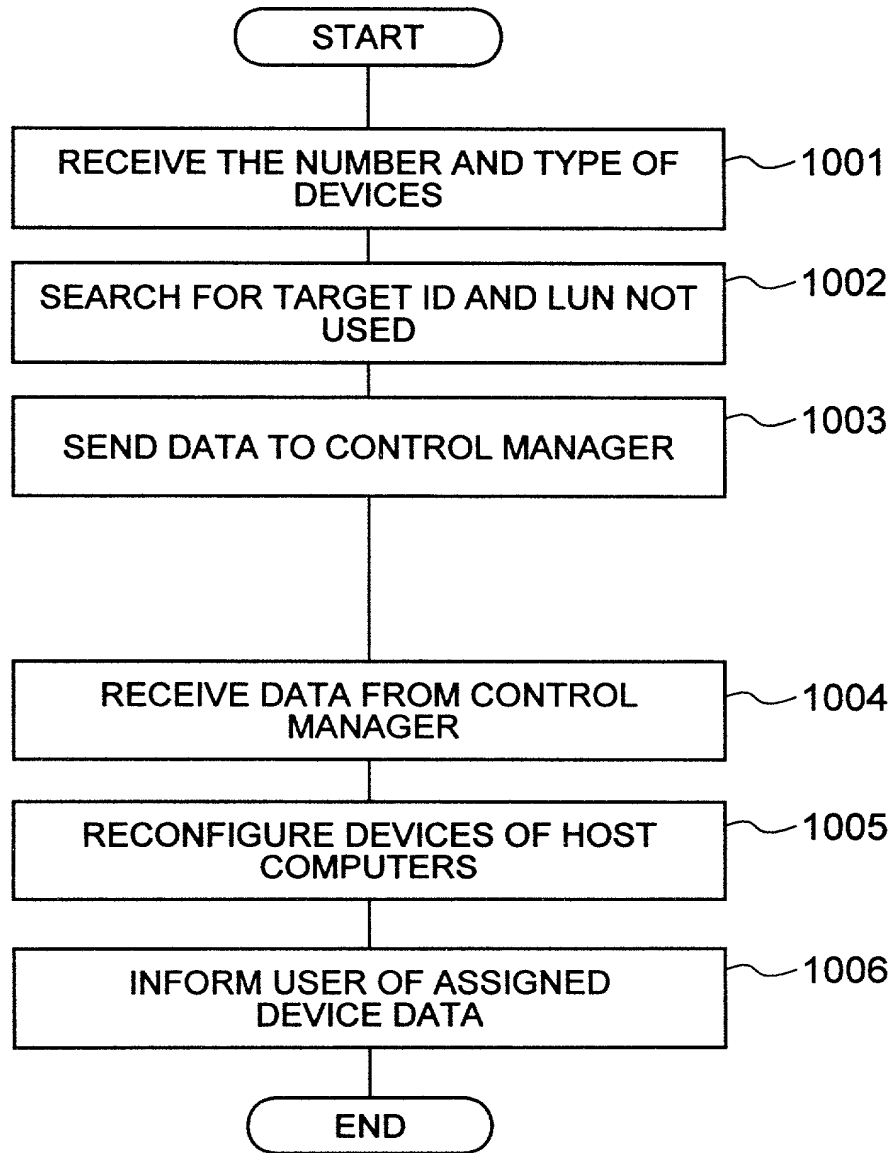


FIG. 5

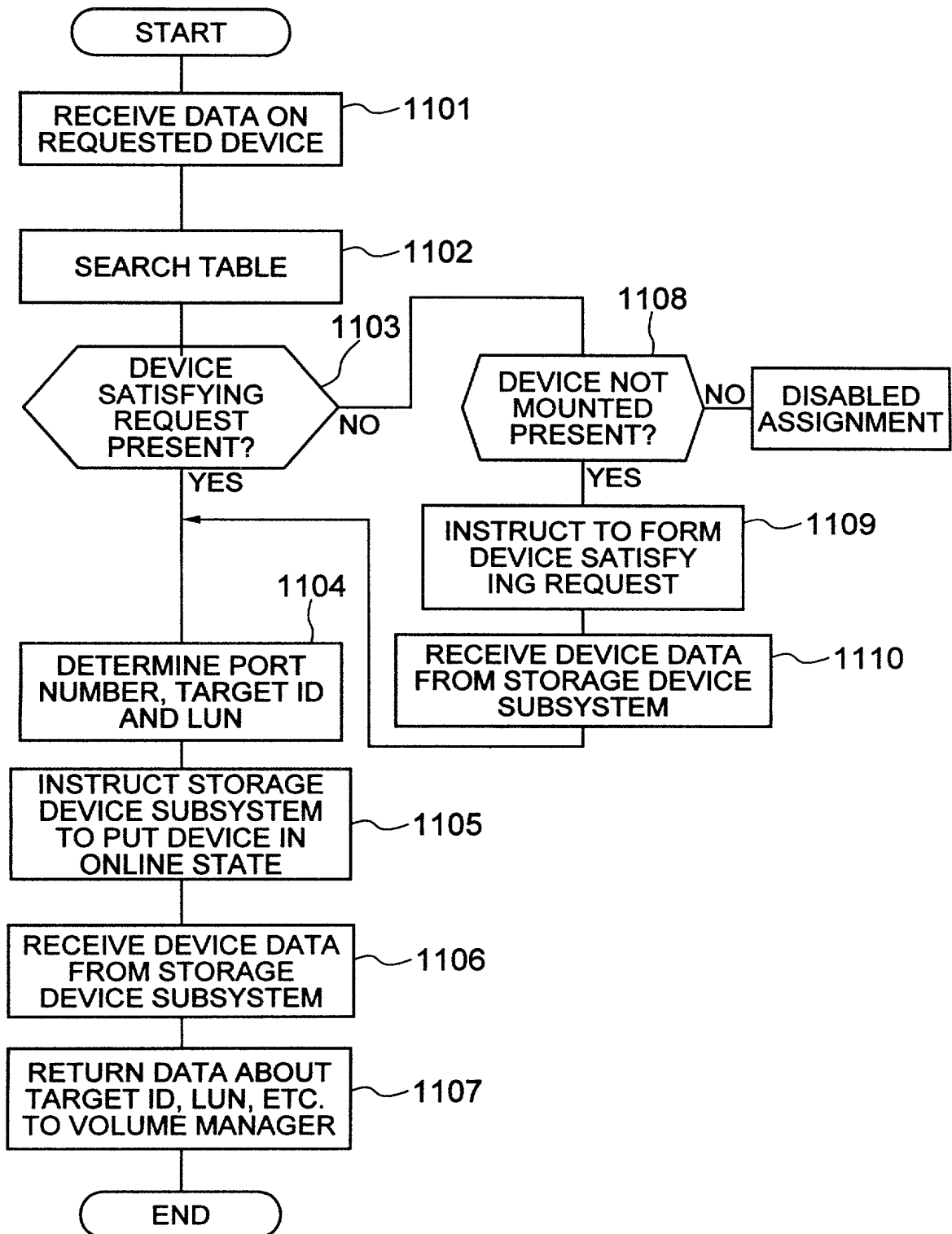


FIG. 6

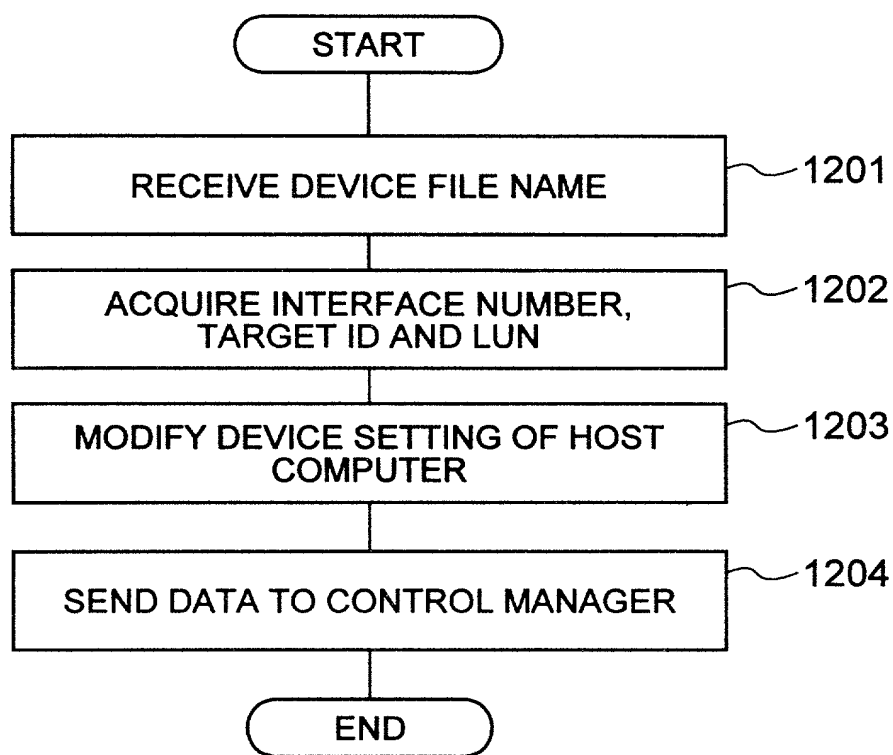


FIG. 7

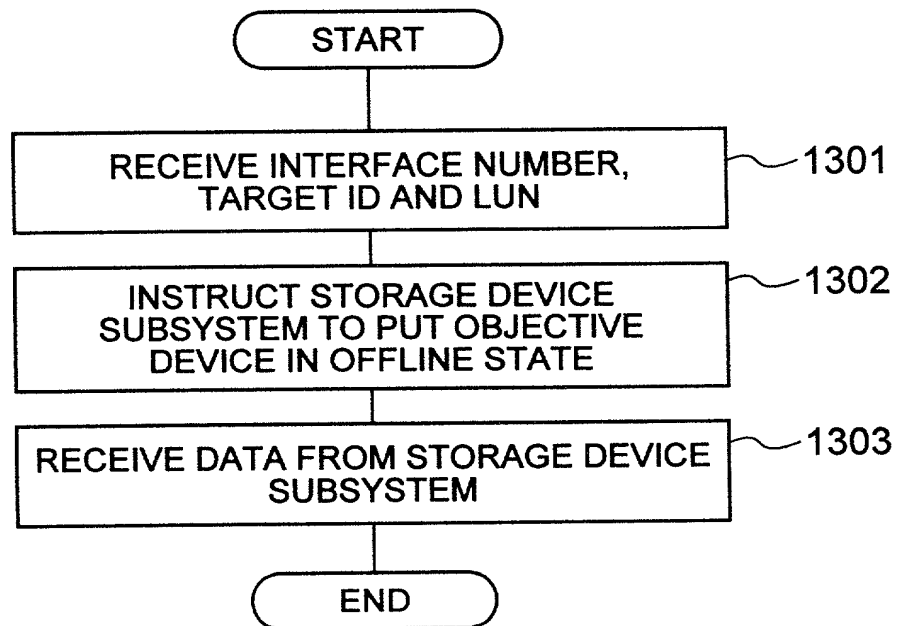


FIG. 8

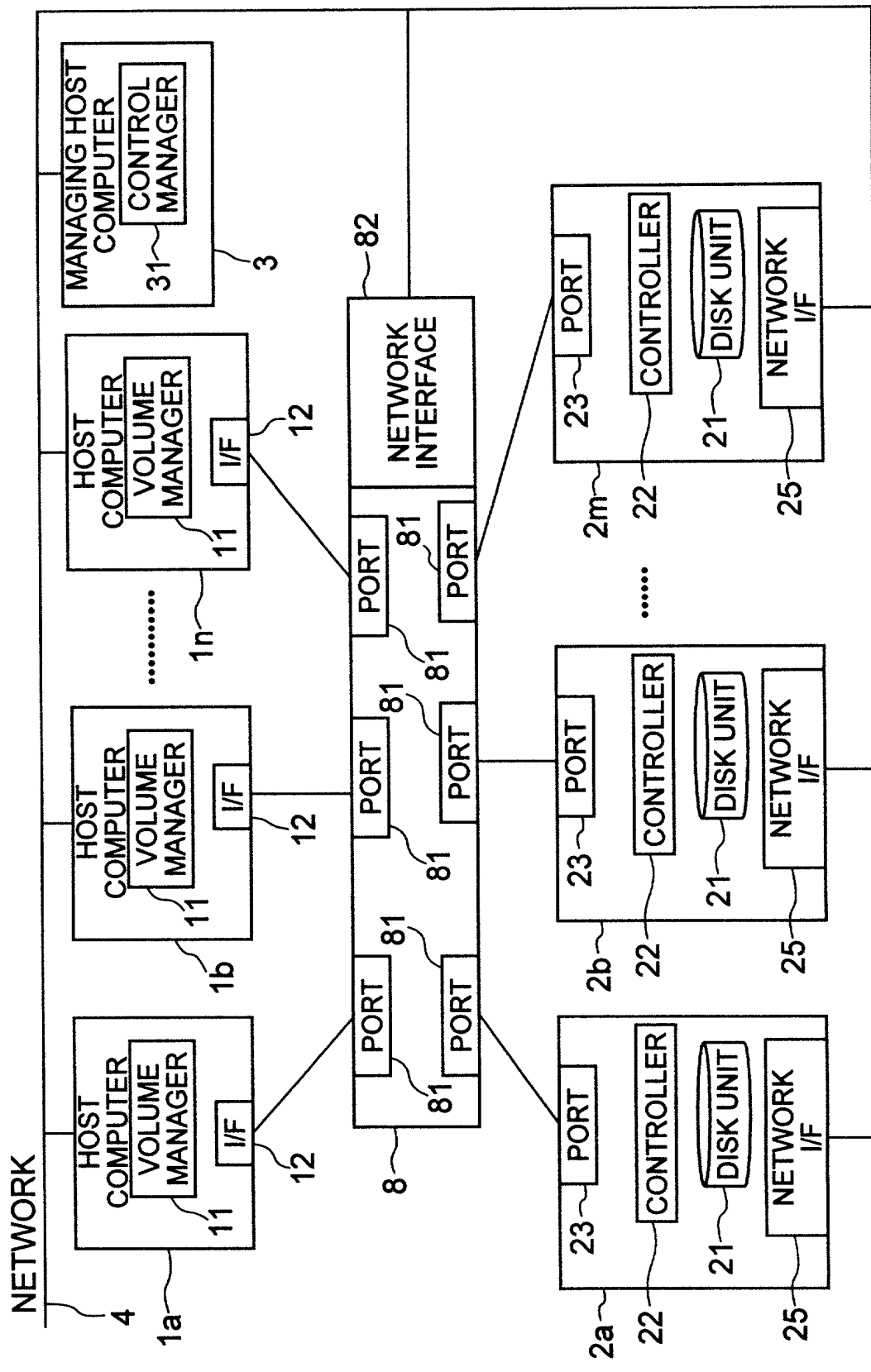


FIG. 9

101 DEVICE NUMBER	102 WWN	103 SIZE	104 CONFIGURATION	105 STATE	106 LUN	107 CONNECTING HOST NAME
0	XXXXXXXX	1000000	RAID 1	ONLINE	0	HOST A
1	XXXXXXXX	1000000	RAID 5	OFFLINE	-	
2	XXXXXXXX	35000	CACHE	ONLINE	2	HOST A,HOST B
.
.
.
.
n	XXXXXXXX	1000000	SINGLE DISK	OFFLINE TROUBLE	0	

FIG. 10

HOST COMPUTER NAME	WWN
HOST 1	XXXX
HOST 2	XXXX
HOST 2	XXXX
⋮	⋮
HOST N	XXX

108

109

FIG. 11

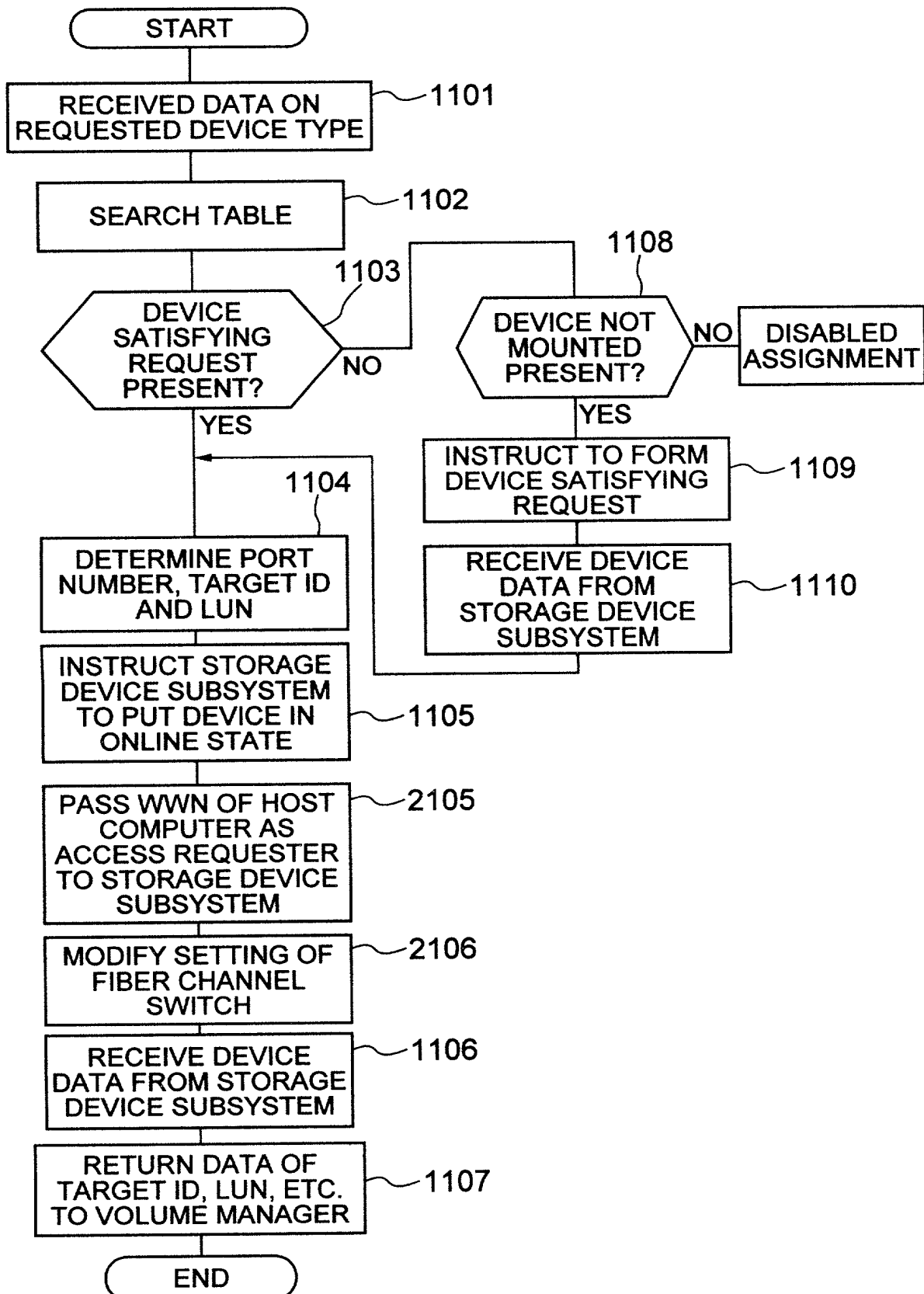


FIG. 12

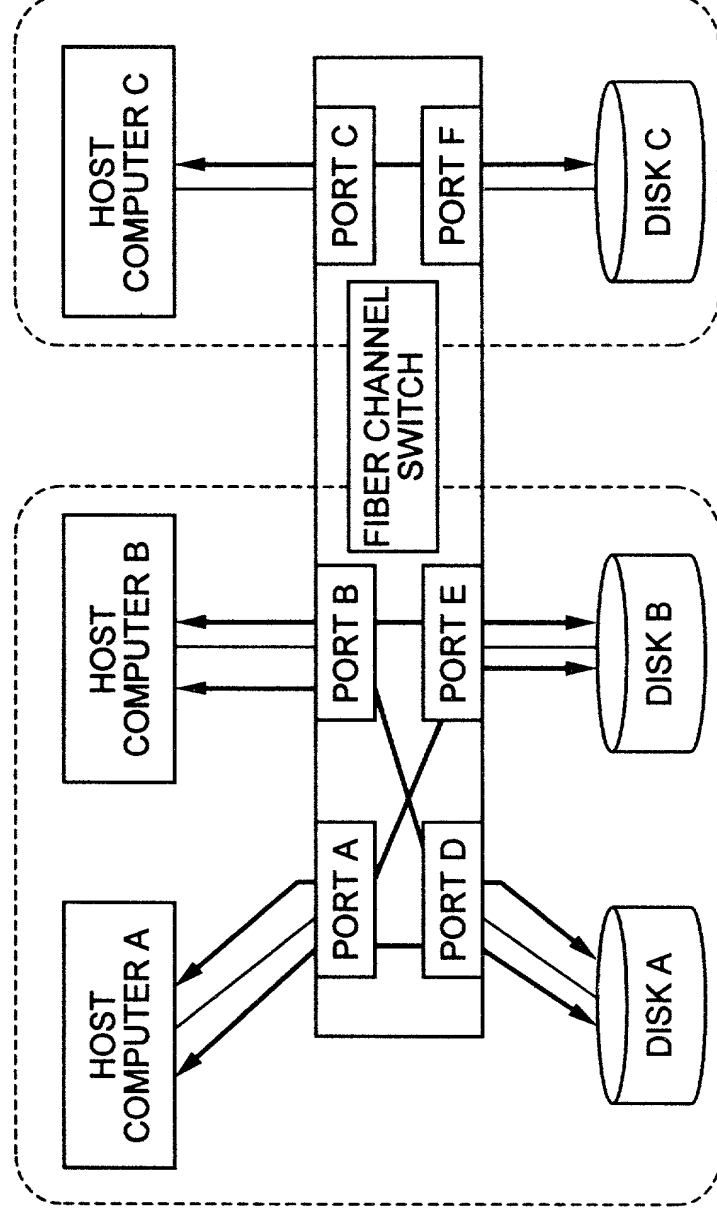


FIG. 14

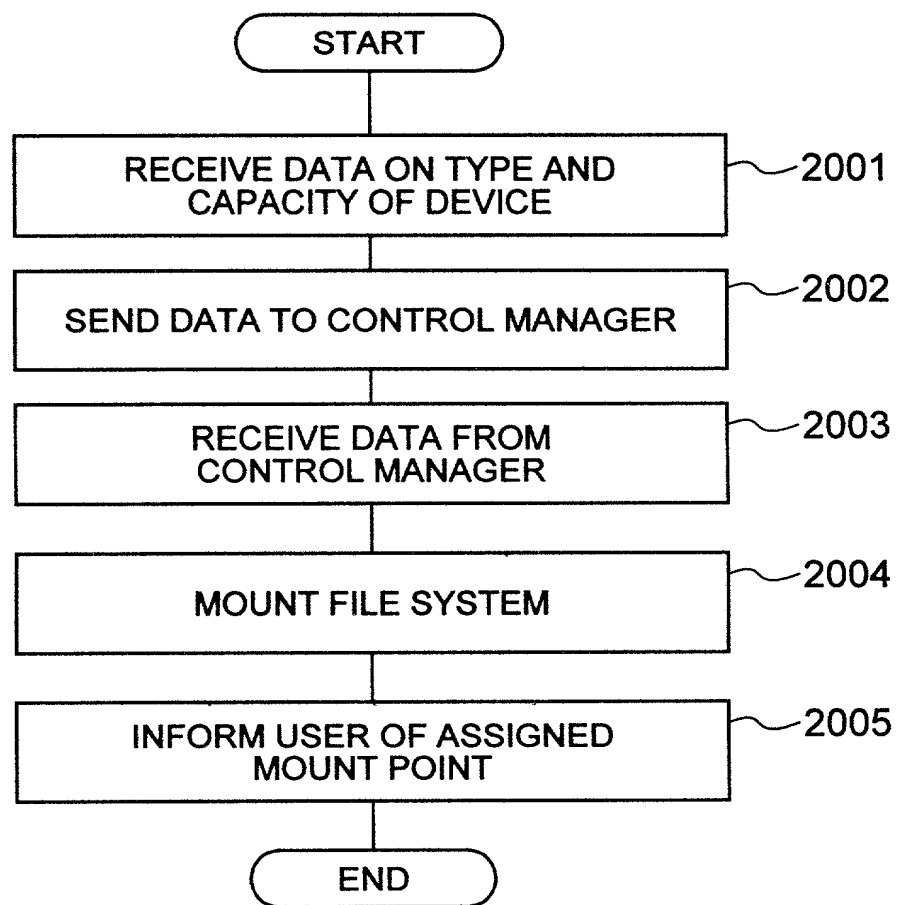


FIG. 15

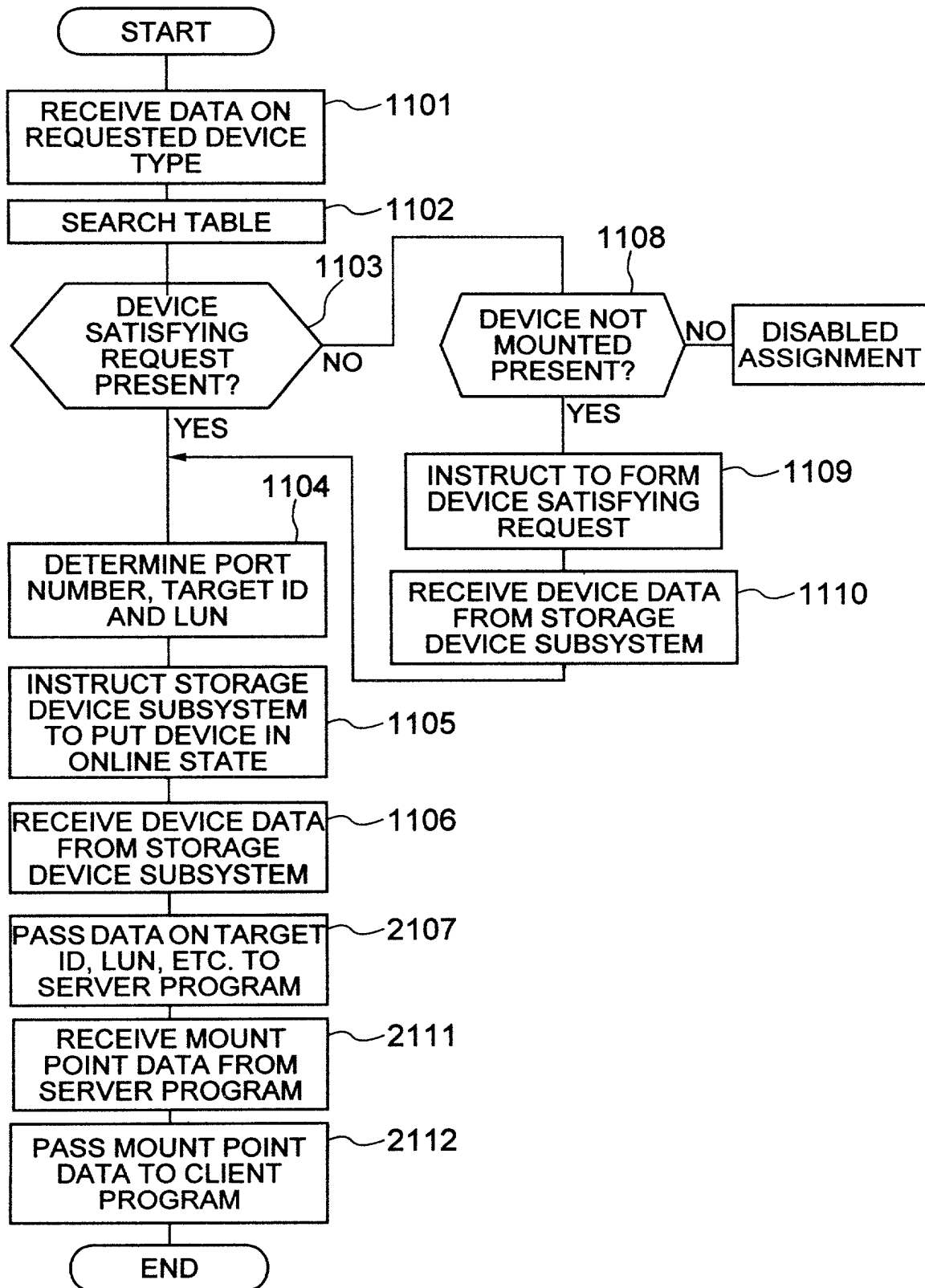


FIG. 16

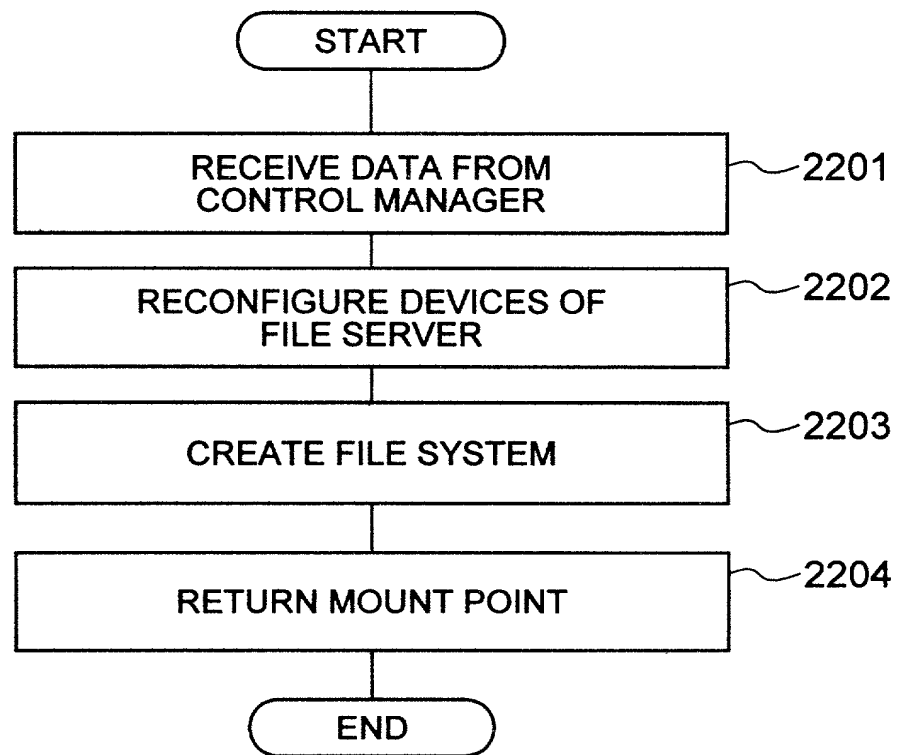


FIG. 17

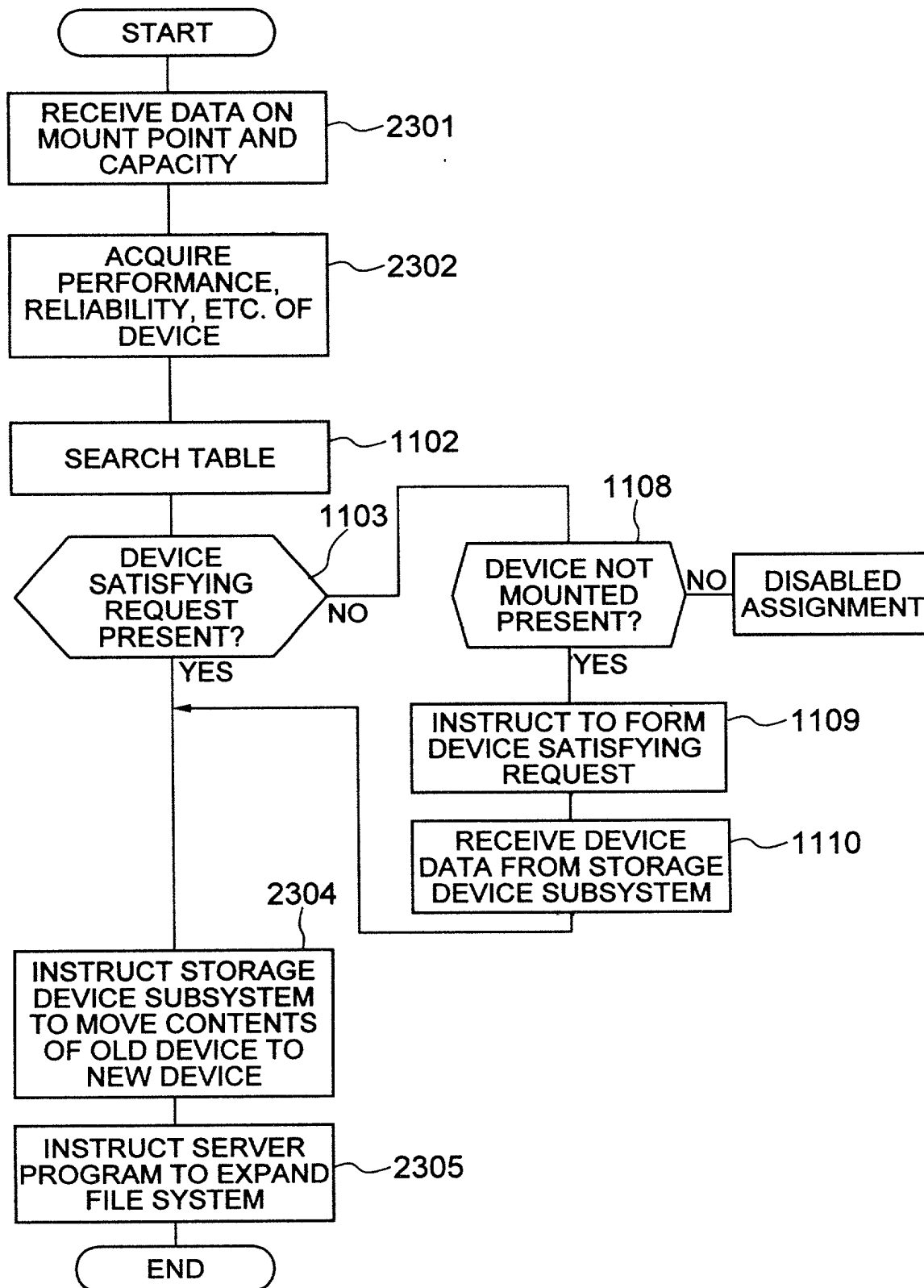


FIG. 18

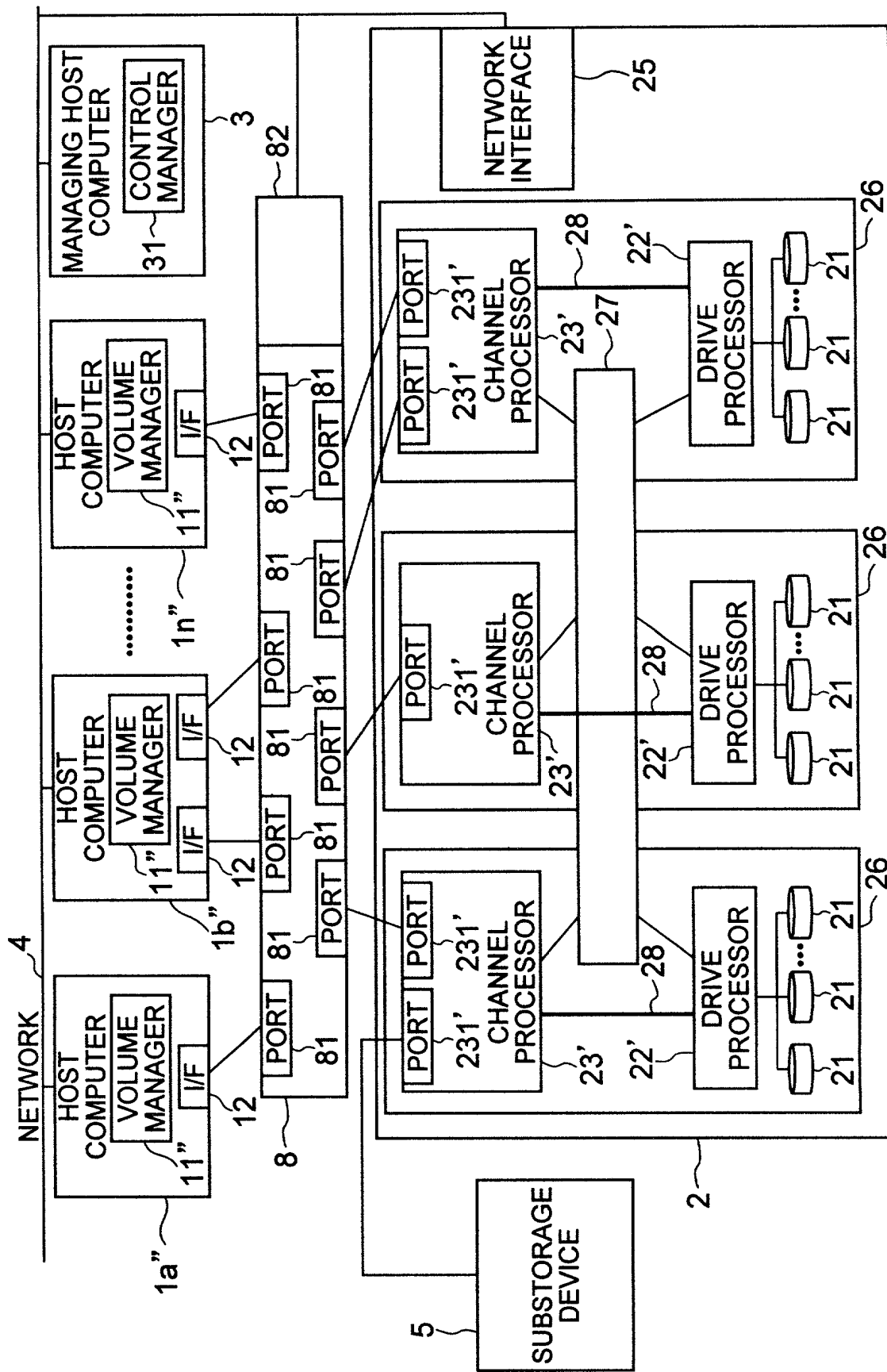


FIG. 19

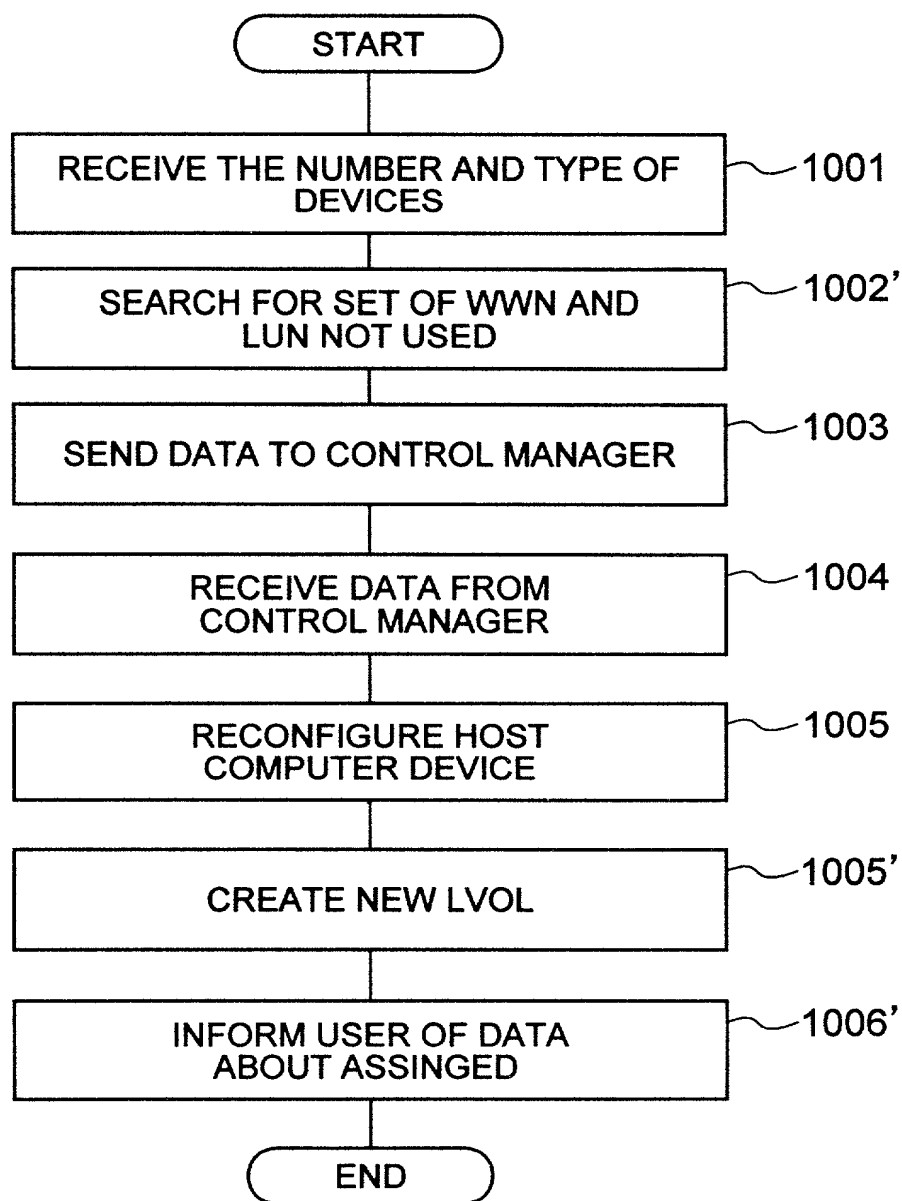


FIG. 20

151	152	153	154	155
LVOL NAME	DEVICE FILE	SIZE	WWN	LUN
lv01	/dev/rdisk/c0t0d0	1000000	XXXX	0
lv02	/dev/rdisk/c0t0d1	500000	XXXX	1
lv03	/dev/rdisk/c1t0d0 /dev/rdisk/c1t0d1	35000 500000	XXXX XXXX	0 1
⋮	⋮	⋮	⋮	⋮
lv0N			XXX	n

FIG. 21

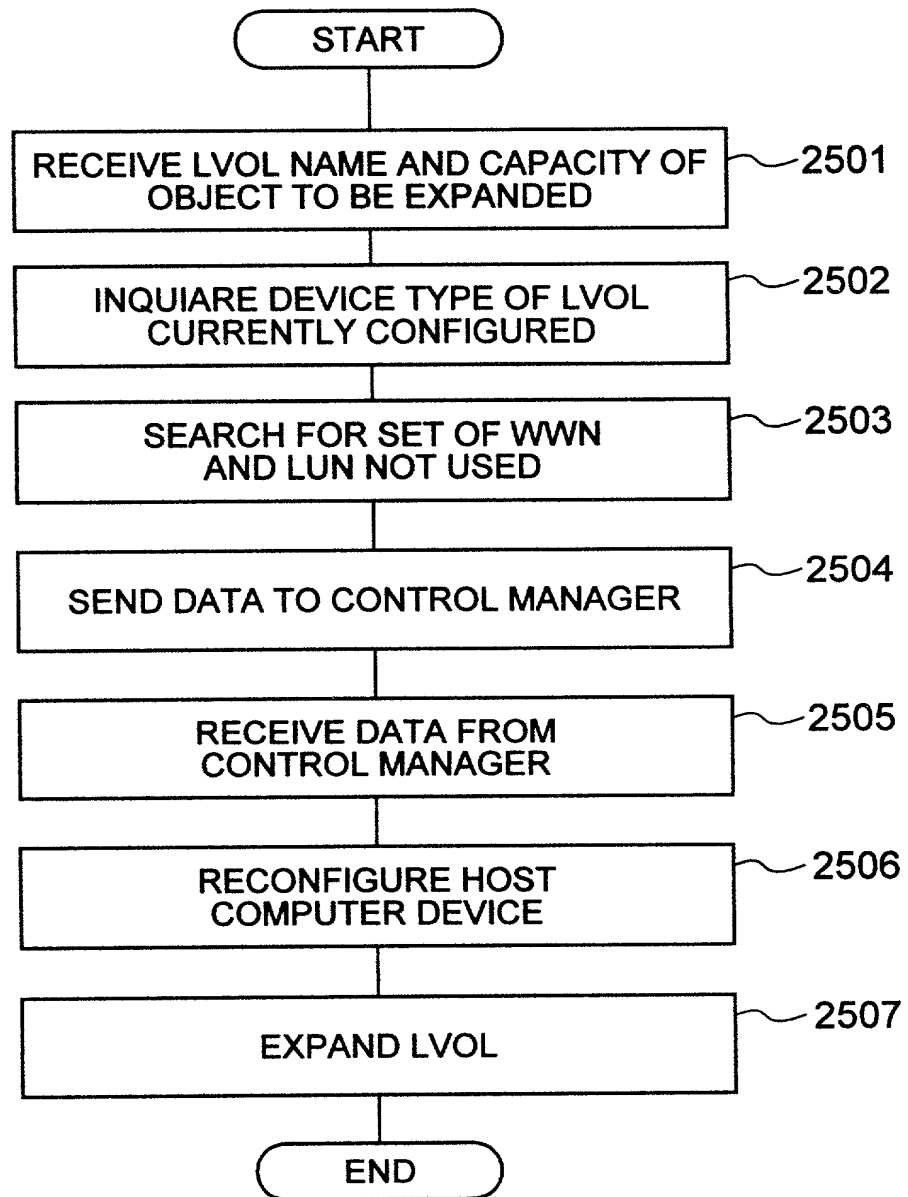
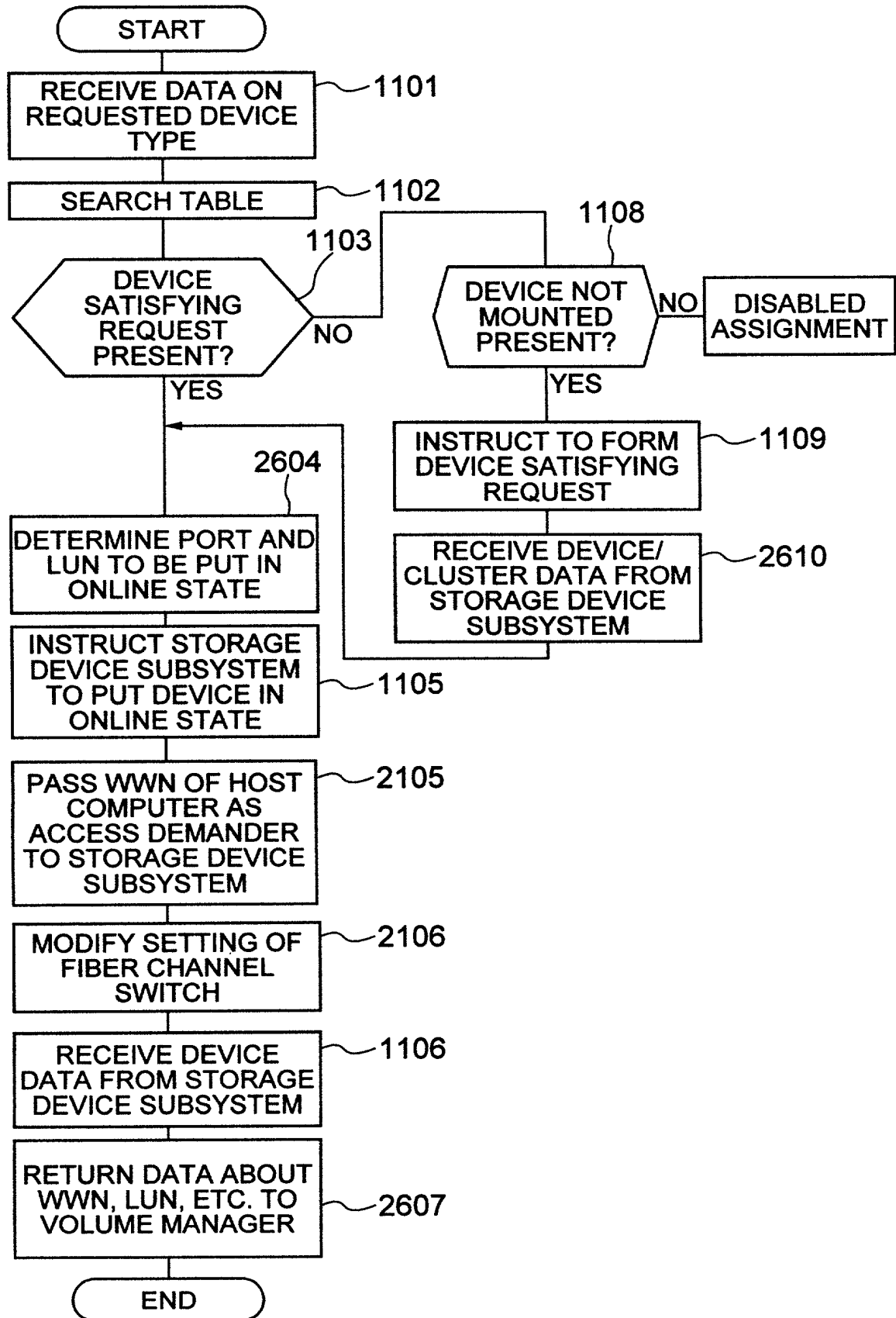


FIG. 22

161 CLUSTER NUMBER	162 PORT NUMBER	163 WWN
1	1 2	XXXX XXXX
2	3	XXXX
3	4 5	
⋮	⋮	⋮
N		XXXX

002280 " 21624360

FIG. 23



ASA-911
E5409-01EK

LIST OF INVENTORS' NAMES AND ADDRESSES

1. Manabu KITAMURA, Tokyo, Japan.
2. Kenji YAMAGAMI, Tokyo, Japan.
3. Tatsuya MURAKAMI, Tokyo, Japan.

002280 " 27824950